

ARI Research Note 97-08

Visualization and Judgmental Forecasting of Simulated Battles

Robert E. Solick, Douglas K. Spiegel,
James W. Lussier, and S. Delane Keene

U.S. Army Research Institute

Fort Leavenworth Research Unit
Stanley M. Halpin, Chief

May 1997

DDIC QUALITY INSPECTED 2

19970825 127



United States Army
Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

**A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel**

EDGAR M. JOHNSON
Director

Technical review by

James Banks

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE 1997, May		2. REPORT TYPE Final		3. DATES COVERED (from... to) April 1991-April 1993	
4. TITLE AND SUBTITLE Visualization and Judgmental Forecasting of Simulated Battles				5a. CONTRACT OR GRANT NUMBER	
				5b. PROGRAM ELEMENT NUMBER 0601102A	
6. AUTHOR(S) Robert E. Solick, Douglas K. Spiegel, James W. Lussier, and S. Delane Keene				5c. PROJECT NUMBER B74F	
				5d. TASK NUMBER 1111	
				5e. WORK UNIT NUMBER H01	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences Fort Leavenworth Research Unit P.O. Box 3407 Fort Leavenworth, KS 66027-0347				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN PERI-RK 5001 Eisenhower Avenue Alexandria, VA 22333-5600				10. MONITOR ACRONYM ARI	
				11. MONITOR REPORT NUMBER Research Note 97-08	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT (<i>Maximum 200 words</i>): <p>Army officers were given information about battles fought in training exercises. They were required to report upon the current situation and to predict future locations and strengths of the forces involved. A battery of cognitive tests was also administered. Accuracy of judgment was associated with experience and with some of the cognitive abilities tested, particularly with memory for spatial relationships, at which the officers excelled. However, the influence of experience was dependent upon the inherent predictability of the scenario—experienced officers did better on a normal mission plan but were less accurate on a plan that was poorly executed. The overall pattern of results suggested that accuracy is strongly influenced by the pace of battle. Static or slowly changing conditions were relatively easy to visualize and predict, but rapidly changing conditions were associated with large increases in error.</p>					
15. SUBJECT TERMS <div style="display: flex; justify-content: space-between; padding: 5px;"> Visualization Forecasting Prediction </div> <div style="display: flex; justify-content: space-between; padding: 5px;"> Memory Spatial Performance measurement </div>					
SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT Unlimited	20. NUMBER OF PAGES 60	21. RESPONSIBLE PERSON (Name and Telephone Number) Robert E. Solick (913) 684-4933
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified			

FOREWORD

The U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Leavenworth Field Unit, conducts a program of research on organizations and leadership in support of the Combined Arms Center. The Field Unit has been involved for several years in research to assess and improve decision making in command groups. Studies have been conducted in the field, in the classroom, and in the laboratory.

The research described in this paper evaluates a technique for measuring the ability of commanders and staff officers to assess a tactical situation and to predict what will happen subsequently. The abilities of the individuals and the characteristics of the situation that are associated with accurate judgment were investigated.

The effort was performed as part of the in-house Basic Research Program.

ZITA M. SIMUTIS
Technical Director

EDGAR M. JOHNSON
Director

VISUALIZATION AND JUDGMENTAL FORECASTING OF SIMULATED BATTLES

EXECUTIVE SUMMARY

Research Requirement:

To develop and test a technique for determining what individual and situational characteristics are associated with visualization of the battlefield.

Procedure:

Officers currently serving in command and staff assignments were given information about battles fought in training exercises. They analyzed the mission plan, observed updates of situation maps, and listened to communications that occurred during force-on-force deliberate-attack missions at the National Training Center. Measures related to situation awareness and to knowledge of future states were scored for accuracy. These were analyzed for their association with demographic factors and with scores from cognitive abilities tests.

Findings:

The overall pattern of results suggests that accurate monitoring and prediction of tactical situations is strongly influenced by the pace of battle. Experience was an advantage in general, but was detrimental when the unit deviated from expected actions. The ability to accurately recall spatial relationships was related to visualization performance.

Utilization of Findings:

The measurement method will be used in subsequent investigations. The spatial memory results and pace of battle hypothesis will be topics of future research.

VISUALIZATION AND JUDGMENTAL FORECASTING OF SIMULATED BATTLES

CONTENTS

	Page
INTRODUCTION	1
METHOD	4
Participants	4
Apparatus	4
Procedure	4
RESULTS	8
DISCUSSION	14
REFERENCES	19
APPENDIX A. BATTLE FORECASTING AND VISUALIZATION--ANSWER BOOK	A-1
B. BATTLEFIELD VISUALIZATION SCENARIOS	B-1

LIST OF TABLES

Table 1. Rank, Duty Position, and Branch of Participants	5
2. Cognitive Testing Results Compared to Reference Values	8
3. MANOVA Results for Performance Variables	11
4. Canonical Correlation Results	12
5. Relationships of Performance Measures with Individual Differences	13
6. Interrelationships Among Performance Measures	13

LIST OF FIGURES

Figure 1. Location Errors in Infantry Task Force Scenario.....	9
2. Location Errors in Armor Task Force Scenario	9
3. Strength Errors in Infantry Task Force Scenario	10
4. Strength Errors in Armor Task force Scenario	10

VISUALIZATION AND JUDGMENTAL FORECASTING OF SIMULATED BATTLES

Introduction

Visualization of the battlefield is an important component in battle command. FM 100-5, Operations (1993) considers it a continuing requirement for commanders. While visualization is not officially defined in Army doctrine, the term visualization generally implies the making of a mental image, usually but not necessary a visual image. Psychological tests of visualization focus on the visual aspects of mental imagery. For example, the Paper Folding Test, which is used in this study, requires the participant to look at a picture of a piece of paper, mentally fold the paper several times, imagine a pencil being poked through the folded paper, mentally unfold the paper, and indicate the arrangement of holes in the now unfolded paper. This visualization, as a psychological ability, involves maintaining and manipulating visual images. Another example of the study of visualization in psychology is work with chessplayers who look at a position on a board and mentally move the pieces, visualizing possible future positions, as well as studies of those who play blindfold chess (Holding, 1985). The military interpretation of the term *visualization* is somewhat broader than the psychological usage. The commander, or staff officer, may receive information from a variety of sources, for example, situation maps, radio nets, logistics charts, as well as from his or her own senses. These sources combine to make an overall understanding of militarily relevant events and their interrelationships. FM 100-5, Operations (1993) implies this meaning of the term:

Visualizing the battlefield is a continuing requirement for commanders. In larger tactical and operational formations, the headquarters normally is the focal point for the flow of information and the resulting planning efforts. Yet commanders of neither large nor small units can visualize the battlefield . . . from a computer screen at the command post (p. 2-14).

The remainder of the paragraph describes the importance of the commanders leaving the command post to assess the state of battle face-to-face as well as aspects of command other than visualization such as influencing the battle, exerting physical and moral presence, and imposing their will to achieve victory.

The activity described above as battlefield visualization is essentially that which is studied by psychologists as mental models. While there is no consistent definition of what exactly a mental model is, one useful working definition is that a mental model is a task and situation-specific mental representation that supports problem-solving and decision-making in a particular context (Zacharias, Illgen, Asdigha, Yara, and Hudlicka, 1995). An important notion in mental model research is that the mental model is constructed from both current information sources and from one's own underlying knowledge base. The creation of mental models goes beyond the psychologist's use of the term visualization; it includes images that are verbal, logical, temporal, spatial, and abstract as well as visual. For example, a commander considering the expenditure of ammunition will probably not imagine a decreasing pile of rounds, and need not use a visual image at all. The notion of battlefield visualization as the creation of a mental model is captured in the phrase "visualize the battle in time, space, and purpose" which arose at the National Training Center and is contained in the following description of visualization by the Battle Command Battle Lab (1994).

Visualization is the act of forming a mental picture of the current and future state based on higher commanders' intent, available information, and intuition. Seeing enemy, friendly, and terrain in terms of time, space, and purpose form the basis of the commander's estimate. While a portion of the desired future state may be dictated by a higher commander's intent, the battle commander must possess the ability to envision his organization's future state within its battlespace.

The concept of visualizing future states implies an important aspect of mental models. Not only do mental models provide a way to organize a collection of related objects, their attributes, and their interrelationships but mental models can also be manipulated or operated upon to produce predictions of future states or outcomes. Thus battlefield visualization implies both understanding what is happening on the battlefield and predicting what will happen. In this paper we use the term *visualization* for the first part, integrating various data sources to form a picture of what is occurring on the battlefield and *forecasting* for the second part, which is making predictions of what will happen.

This report describes a procedure which was developed to measure battlefield visualization and forecasting. Army officers were given information from battles that had been fought at the National Training Center. They reviewed the plan and judged the probable success of the operation, then received information about how the battle was progressing (from the command radio net). They were required to report upon the current situation, giving a narrative account, estimating unit locations, and estimating current strength of friendly and enemy forces. They were also asked to discuss what would happen during the next hour of battle and to estimate future unit locations and strengths. Each officer reviewed one battle, pausing frequently to give judgments on current and future elements. The essence of the measurement procedure is to ask the participants to provide an evaluation of a tactical situation that can be verified by comparing it to what actually happened during a training exercise. Validation of the procedure requires demonstrating that responses vary in a manner consistent with variations in the conditions presented and with variations in the capabilities of the people making the judgments. The empirical issues of interest initially are answers to the questions: What makes the task difficult? What makes someone good at it? In addition, each experimental session included a battery of tests of basic cognitive skills and an attempt was made to relate visualization and forecasting performance to these basic cognitive skills.

The National Training Center at Fort Irwin, California provides force-on-force and live fire training for armor and mechanized infantry units (Bolger, 1986). As part of the force-on-force training, tactical operations are conducted by a U.S. battalion task force against a simulated Warsaw Pact type of opposing force (OPFOR). Both sides have low power lasers attached to their weapons systems which emit a coded pulse when a blank round of ammunition is fired. Soldiers and vehicles are equipped with sensing devices which detect the laser impulses and determine whether the person or object has been "killed" or "disabled". Vehicle locations and firing events are broadcast to a data-logging facility. Message traffic from a variety of tactical radio networks is also recorded. All of this information, plus the written orders and map overlays developed by the staff and the training feedback sessions and final exercise report developed by the trainers, is stored in the National Training Center archive. The archive is adequate to recreate

most of the conditions and actions during an exercise as well as much of the information available to the staff of the unit involved. The major exceptions are information gleaned through personal observation, planning details worked out in meetings and rehearsals that do not appear in the formal order, and information that was transmitted by means other than radio.

The objectives of this project were to validate the measurement procedure, to obtain information on individual characteristics related to tactical judgment, to look for useful empirical generalizations, and to relate the results to cognitive theory. The measurement procedure was designed to examine substantive tactical judgments under controlled conditions. It provides an objective standard of accuracy, while replicating enough of the context of battlefield decisions to permit generalization of the results to operational conditions. Two judgments that must be arrived at repeatedly during a tactical operation are "What is the current situation?" and, "What will happen next?" The answers to these questions are of considerable practical importance; hence, the emphasis on the individual characteristics associated with good judgment and the concern with the context in which the judgments are solicited. Characteristics of the situation that make accuracy difficult to achieve may be countered through procedural or technological changes. Similarly, an acceptable theoretical account of how information is evaluated, integrated into internal representations of tactical situations, and operated upon to produce situation awareness and expectations of future events, may suggest better approaches to training or assisting staff officers.

A variety of cognitive theories was considered suitable for explaining one aspect or another of visualization and forecasting performance. For example "intuitive statistician" approaches (Kahneman, Slovic, & Tversky, 1982), which base predictions upon knowledge of relative frequencies, can account for quantitative estimates in many situations, while verbal accounts of a current situation or a future state are better handled by cognitive schema concepts such as "scripts" or "plans" (Beach, 1990; Schank and Abelson, 1977). Similarly, episodic models and causal mental models (Johnson-Laird, 1983) may be invoked to handle the variety of strategies that an individual may select to meet the demands of an experimental situation. Constraint-based reasoning (Thuring and Jungerman, 1986) incorporates a subset of those strategies which are particularly apt for the rapid generation of judgments under uncertainty. In advance of the data, theoretical considerations played a role in selection of variables and response methods, but did not lend themselves to prediction of experimental outcomes.

Method

Participants

Sixty-six officers at eight Army installations participated in the study during Umbrella Weeks¹ in the first seven months of 1992. Table 1 shows their ranks, duty positions, and branches. This was a convenience sample. Battalion Commanders, Intelligence Officers (S2s), and Operations Training Officers (S3s) were requested and, indeed, formed the large majority of our participants. Others were substituted when necessary and possible. Half of these officers were put through the protocol for each of the two experimental scenarios, which will be described presently. They were divided between the two scenarios as follows: LTC 8/8 (i.e., eight in the first scenario and eight in the second), MAJ 9/11, CPT 13/9, 1LT 3/3, and 2LT 0/2. The mean age was 34 years; mean Army service was 11 years; and mean time in a heavy (Armor or Mechanized) unit was 54 months.

Apparatus

The National Training Center data base, housed at the Army Research Institute Field Unit at the Presidio of Monterey, was scanned for instances of Field Training Exercises (FTXs) for which there were enough data (graphic, auditory, and electronic) that we could reconstruct a close approximation to the scenario of the FTX. We found two suitable exercises and made copies of the map overlays, command net communications, and digital data (of team locations, strengths, and player status). The audio tapes of the command net communications were transcribed and edited to eliminate much of the noise and dead air time. These were divided into segments covering 15, 30, and 60 minutes of the battle but the actual times required to play the tape segments were approximately 5, 10, and 20 minutes. The transcripts were tagged with the real times.

Procedure

The Gestalt Completion, Myers-Briggs, and Hidden Figures tests were administered to the first twelve subjects. For the remainder the Myers-Briggs was dropped and in its place Paper Folding, Building Memory, and Advanced Vocabulary were substituted. Copies of these tests were put together into booklets, one for each participant. Three pages preceded the tests in this booklet. The first page described the purpose of the research and the role of the participant. The second page was an informed consent form which we asked the participant to sign. The third page requested name (optional), age, rank, branch, years of commissioned service, current duty position, number of NTC rotations, year of most recent NTC experience, which (if any) of the following schools were completed: Combined Arms and Services Staff School (CAS³),

¹ Each Army division in the continental United States sets aside one week each year during which its personnel are available to take part in research projects. These are called Umbrella Weeks.

Table 1

Rank, Duty Position, and Branch of Participants

Rank	Duty Position	Branch					Total
		IN	MI	AR	AV	FA	
LTC	Battalion Commander	6		6	2	2	16
MAJ	Battalion S3	11		5	2	1	19
	Battalion Exec. Officer	1					1
CPT	Assistant Brigade S2		1				1
	Battalion S2		11		2	2	15
	Battalion S3	3					3
	Company Commander	1					1
	Headquarters and	1					1
	Headquarters Company						
	Commander						
	Research & Dev COORD			1			1
1LT	Assistant S2		4				4
	Assistant S3	1					1
	Battalion S2		1				1
2LT	Assistant S2		1				1
	Assistant S3	1					1
TOTAL		25	18	12	6	5	66

Command and General Staff College (CGSC), Tactical Commanders Development Course (TCDC), War College, and number of months of experience in an armored or a mechanized unit. The last item requested was not printed on the page, due to an oversight, but was orally requested and the response was written on this page. Two more booklets, one for each scenario, were produced and duplicated. These booklets contained instructions and response pages for each stage of the protocol. They were similar but not identical for the two scenarios. (See Appendix A for a copy of the Scenario 6 booklet. In our original search for scenarios we numbered them. Numbers four and six were the ones we eventually used.)

Myers and McCaulley (1988) describe the Myers-Briggs as follows: "The aim of the MBTI is to identify . . . the basic preferences of people in regard to perception and judgement, so that the effects of each preference . . . can be established by research and put to practical use."

Descriptions of the tests, as given by Ekstrom, et al. (1976) are as follows:

Gestalt Completion - "Drawings are presented which are composed of black blotches representing parts of the objects being portrayed. The subject writes down the name of the objects, being as specific about them as possible."

Hidden Figures - "The task is to decide which of 5 geometrical figures is embedded in a complex pattern. The difficulty level of this test is high. This particular form of the test was modified from an earlier test developed in connection with a project designed to study field independence."

Paper Folding - "For each item successive drawings illustrate two or three folds made in a square sheet of paper. The final drawing of the folded paper shows where a hole is punched in it. The subject selects one of 5 drawings to show how the punched sheet would appear when fully reopened."

Building Memory - "The subject is asked to indicate the location of a number of buildings seen on a previously studied map."

Advanced Vocabulary - "This is a 5-choice synonym test consisting mainly of difficult items."

The following list gives the procedural sequence used for collecting data. A review of Appendix A will help understanding of this list. In addition, Appendix B gives the brief protocols used by the experimenters and will probably help the reader's comprehension.

1. Read the Plan
2. Assess the plan
3. Predict events for the next hour
 - 3.1. Example map for response training
 - 3.2. Predict team locations at 15, 30 and 60 minutes hence
 - 3.2.1. Recorded on map in response booklet
 - 3.3. Estimate percent and predict future (15, 30, and 60 min.) strengths
 - 3.3.1. Recorded on form in response booklet
4. Listen to commo tape, segment #1, 15 min.
 - 4.1. Discuss understanding obtained from commo tape
 - 4.1.1. Tape recorded
 - 4.2. Estimate team locations on map
 - 4.2.1. Recorded on map in response booklet
 - 4.3. Estimate friendly and enemy strengths, give confidence
 - 4.3.1. Recorded on form in response booklet
 - 4.4. Predict events for the next hours
 - 4.4.1. Tape recorded
 - 4.5. Predict team locations at 15, 30, and 60 minutes hence
 - 4.5.1. Recorded on map in response booklet
 - 4.6. Estimate team locations on map
 - 4.6.1. Recorded on map in response booklet
 - 4.7. Estimate present and predict future (15, 30, and 60 min.) strengths

- 4.7.1. Recorded on form in response booklet
- 5. Listen to commo tape, segment #2, 30 min.
 - 5.1. Discuss understanding obtained from commo tape
 - 5.1.1. Tape recorded
 - 5.2. Estimate team locations on map
 - 5.2.1. Recorded on map in response booklet
 - 5.3. Estimate friendly and enemy strengths, give confidence
 - 5.3.1. Recorded on form in response booklet
 - 5.4. Predict events for the next hour
 - 5.4.1. Tape recorded
 - 5.5. Predict team locations at 15, 30, and 60 minutes hence
 - 5.6. Estimate team locations on map
 - 5.6.1. Recorded on map in response booklet
 - 5.7. Estimate present and predict future (15, 30, and 60 min.) strengths
 - 5.7.1. Recorded on form in response booklet
- 6. Listen to commo tape, segment #3, 60 min.
 - 6.1. Discuss understanding obtained from commo tape
 - 6.1.1. Tape recorded
 - 6.2. Estimate team locations on map
 - 6.2.1. Recorded on map in response booklet
 - 6.3. Estimate friendly and enemy strengths, give confidence
 - 6.3.1. Recorded on form in response booklet
 - 6.4. Predict remainder of the battle
 - 6.4.1. Tape recorded
 - 6.5. Estimate probability of Blue Forces success
 - 6.5.1. Written in response booklet

Coordinates of estimates and predictions of team locations were obtained from the maps upon which judgments of current and future locations were marked by participants. The judgment requested was for the participant to mark the position of the lead element of each team. Undirected distances from marked locations to true locations were computed for each Company Team in a scenario (four teams in the armor battalion scenario and five teams in the mechanized infantry scenario). For each individual, geometric means of the distances were computed across company teams. The resulting statistics were used as location accuracy measures in all subsequent analyses. Strength data were handled in a comparable fashion: True values were subtracted from estimated values separately for enemy and friendly tanks and armored personnel carriers. The absolute values of the errors were added across equipment types separately for friendly and enemy forces to form the basic statistics for subsequent analyses of strength data.

Results

The individual cognitive tests were scored according to the procedures described in the Manual for Kit of Factor Referenced Cognitive Tests (Ekstrom, French, Harman and Dermen, 1976). Univariate results are shown in Table 2, compared to reference values found in other studies, from Ekstrom, et al (1976), pp 11-15.

Table 2

Cognitive Testing Results Compared to Reference Values

Test	Experimental Population		Reference Values		Reference Sample
	Mean	S.D.	Mean	S.D.	
Hidden Figures	12.8	5.7	14.0	6.4	167 college males
Gestalt Completion	14.1	2.8	13.8	--	215 college students
Building Memory	18.4	4.4	10.9	4.7	542-574 male Naval recruits
Advanced Vocabulary I	18.6	4.7	13.8	5.2	181 college students
			13.3	--	83 Army enlistees
Paper Folding	10.5	3.9	10.4	--	82 Army enlistees

Our sample of Army officers, all of whom are college graduates, some with advanced degrees, gave results comparable to the reference populations, except that they were substantially higher on Advanced Vocabulary and Building Memory. Building Memory is a test of memory for the spatial location of symbols representing buildings on a sketch map. The maximum score that can be obtained is 24. Six of the 54 Army officers taking this test achieved the maximum without any apparent effort, while only 3 scores were lower than the mean for the reference group. The test was administered with a four minute study period and a four minute recall period for each sketch map. Most of our participants used less than half of the time allotted and waited patiently, or otherwise, for the remaining time to expire.

Figures 1 and 2 illustrate the average location errors for the 33 participants viewing the mechanized infantry and tank battalion scenarios, respectively. Figures 3 and 4 show the data for strength judgments, combining errors for friendly and enemy forces. There is a clear trend toward larger errors in prediction with increasing time span of the projection, as one would expect from compounding of errors (the 15 minute projection is from a known current location, the 30 minute projection from the assumed 15 minute location, etc.). This holds for both strength and location results. It might be assumed that increasing familiarity with the scenario and the task, as would occur across iterations, would lead to more accurate judgments. However, this was not

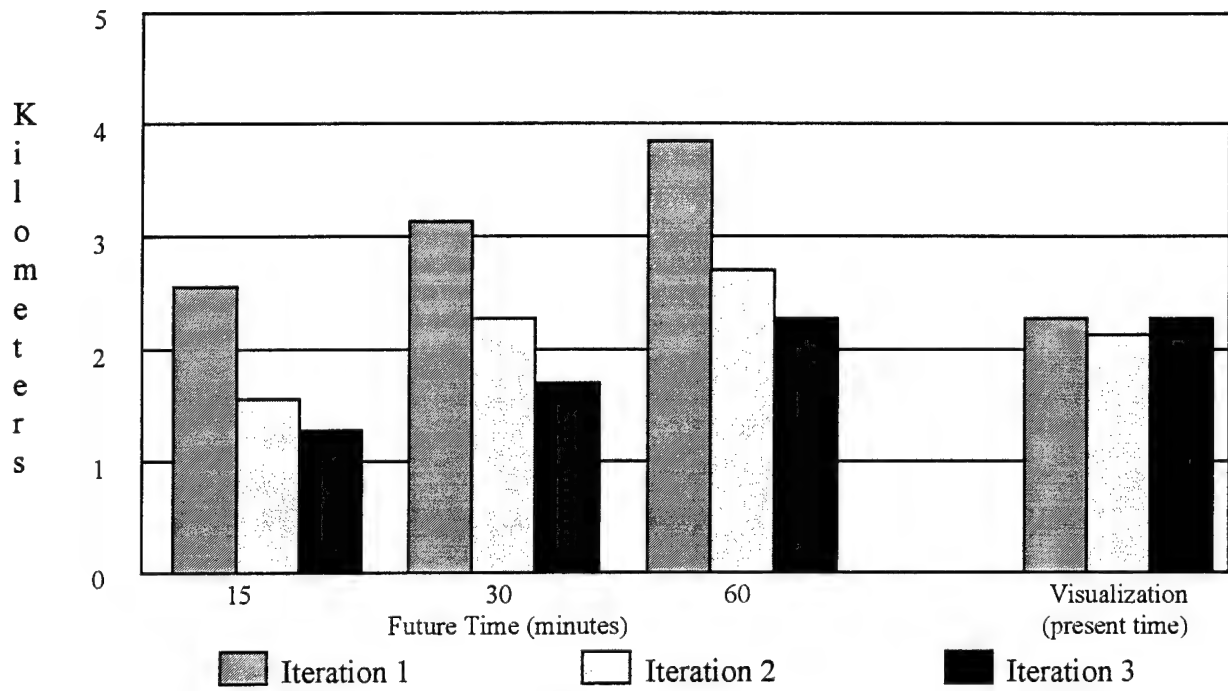


Figure 1. Location Errors in Infantry Task Force Scenario.

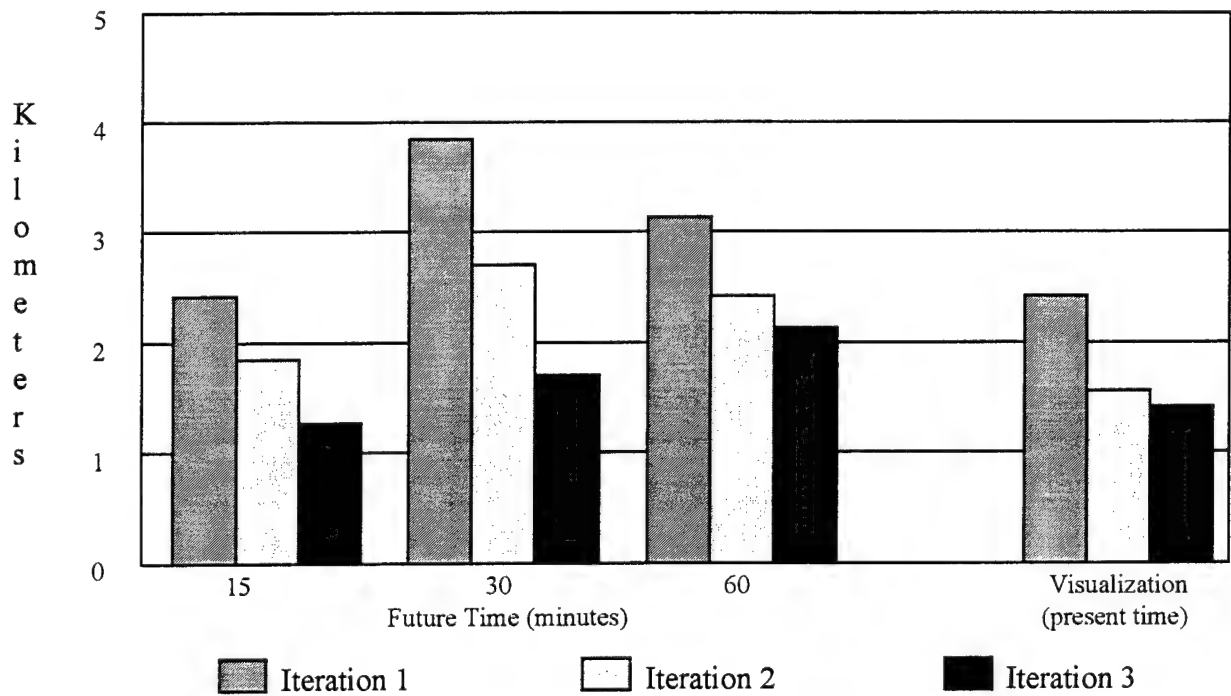


Figure 2. Location Errors in Armor Task Force Scenario

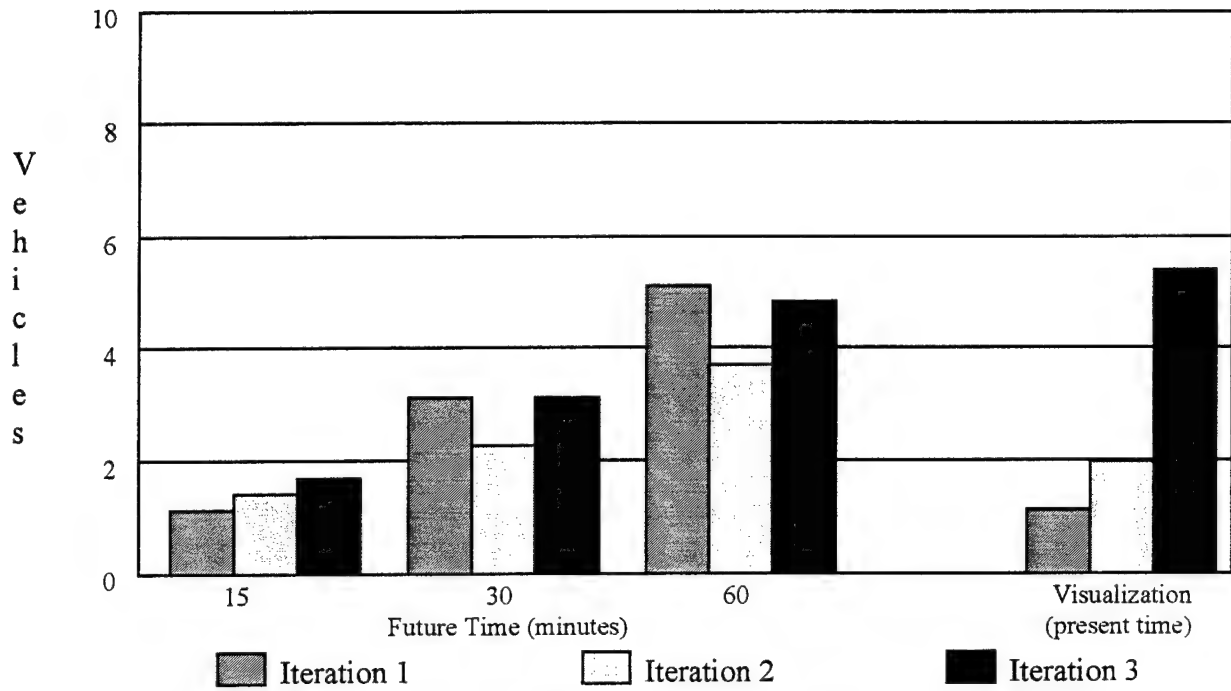


Figure 3. Strength Errors in Infantry Task Force Scenario.

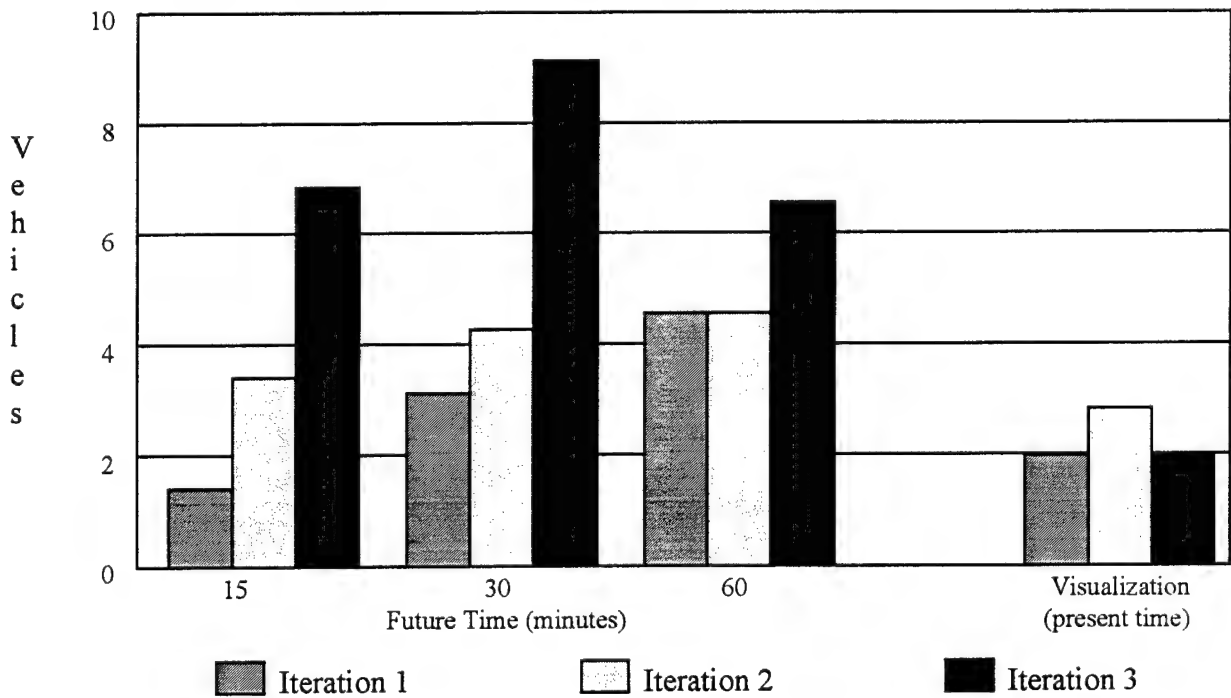


Figure 4. Strength Errors in Armor Task Force Scenario

task, as would occur across iterations, would lead to more accurate judgments. However, this was not generally true. Predicted locations tended to be more accurate across iterations in both scenarios, while visualized locations tended toward increased accuracy only for the armor scenario. Predicted strengths tended to be less accurate across iterations, especially in the armor scenario, while visualized strengths became much less accurate in the infantry scenario, but not in the armor scenario. A simple way of characterizing these apparently complex results will be discussed presently. A multivariate analysis of variance was performed to confirm that the visually obvious results were, in fact, statistically significant. Table 3 presents these results, along with a consideration of the covariate military rank (which correlates very highly with age and years of experience, and is closely related to active duty job).

Table 3

MANOVA Results for Performance Variables

Factor	Pred Loc	Pred Str	Vis Loc	Vis Str
Scenario		***	**	***
Rank	***		*	**
Scenario X Rank				
Iteration	***	***	***	***
Iteration X Scenario		***	***	***
Iteration X Rank		*		
Iter X Scen X Rank	*			
Time	***	***		
Time X Scenario	***	***		
Time X Rank	***			
Time X Scen X Rank				
Iteration X Time	*	***		
Iter X Time X Scen	***	***		
Iter X Time X Rank	***	*		
Iter X Time X Scen X Rank				
*p <.05	**p <.01	***p <.001		

A Canonical Correlation Analysis (Table 4) showed that approximately two thirds of the variance in a weighted linear combination of the accuracy measures for predicting and visualizing strengths and locations could be accounted for by a weighted linear combination of measures of individual differences, including the five cognitive test scores, age, and months of experience in armor or mechanized infantry units. Scenario was included in the set of predictors because of the interaction of this factor with the accuracy measures.

Table 4

Canonical Correlation Results

1st Solution ($p < .0001$), Other Solutions $p > .05$

Canonical Correlation	Adjusted Can Corr	Approx St. Error	Squared Can Corr
0.82	0.78	0.05	0.67
Standard Canonical Coefficients			
<u>Individual difference measures</u>		<u>Accuracy measures</u>	
Scenario	0.79	Predict Location Error	-0.19
Gestalt	-0.02	Predict Strength Error	0.68
Hidden Figures	-0.04	Visualize Location	-0.21
Paper Folding	0.32	Visualize Strength	-0.43
Building Memory	0.11		
Adv. Vocabulary	-0.03		
Years of Age	0.25		
Months in Heavy Unit	-0.09		

As shown in Table 5, none of the bivariate associations are particularly strong, in fact, the highest correlation, between Paper Folding and predict strength, is in the "wrong" direction (better test performance is associated with less accurate predictions). The results for Building Memory are potentially interesting, especially given the restricted range of very high scores achieved by the participants. Both the cognitive test and the experimental visualization tasks require one to remember and interpret relationships, either those viewed on a map or information derived from the audio tape of the command net.

Table 5

Relationships of Performance Measures with Individual Differences

	Error Scores			
	Predict Locations	Predict Strengths	Visualize Locations	Visualize Strengths
Gestalt	.05	-.17	-.21	.13
Hidden Figures	-.11	-.14	-.16	-.20
Paper Folding	-.18	.43*	-.21	-.31*
Building Memory	.09	.27	-.32*	-.29*
Adv. Vocabulary	-.28*	-.10	-.28*	-.10
Age	-.27*	.16	-.28*	-.14
Months in Heavy Und	.12	.06	-.15	-.11

* $p < .05$ Pearson r , n ranges from 54 to 66

A final set of relationships is presented in Table 6, showing associations among the measures of accuracy. There may be a relationship between the ability to visualize current locations and predict future locations. Otherwise, the ability to perform any of the experimental tasks well does not reveal much about how one would perform on the other tasks.

Table 6

Interrelationships Among Performance Measures

	Predict Strengths	Visualize Locations	Visualize Strengths
Predict Locations	.05	.31*	.21
Predict Strengths		-.17	-.12
Visualize Locations			.24

* $p < .05$ Pearson r , $n = 65$ or 66

Discussion

The first objective of this project was to demonstrate the measurement technique. The results indicate that the technique was effective. Evaluations of militarily significant features of the situation were obtained from the participants. Accuracy varied in reasonable ways with the capabilities and experience of the individuals and with characteristics of the situation. The response measures were sensitive to scenario differences, to variations within scenarios, and to variations in the response procedure.

The next two objectives were to identify empirical generalizations which would help to characterize what makes a judgment difficult and what makes an individual good at it.

To take the second point first, there is not a simple relationship between military training or experience and accurate performance on our task. While there is a slight statistical advantage for older, more experienced, higher ranking officers, some of the best individual results were supplied by young Captains and one 1st LT, and some of the poorest individual efforts were turned in by more experienced officers. Experience tended to help on the location judgments, but to hinder accuracy on strengths judgments.

The influence of experience (as indicated by military rank) differed for the two scenarios. This might be accounted for on the basis of the inherent predictability of the scenarios. A normal plan, implemented in a straight forward manner, should lead to more predictable results and more readily visualizable intermediate states than an eccentric plan or one that is badly implemented. In the latter case, the more you know about what actions should be taken and what results should be achieved, the less likely you are to deduce the events that actually occurred. Our mechanized infantry scenario was of this type. As one disgruntled battalion commander put it: "How do you expect me to know what they are doing when they don't know themselves?"

A second factor minimizing the influence of experience is restricted range. While years of military service varied from one to twenty-two, only two of the participants could be characterized as novices. The rest had dealt with historical, simulated, and in some cases, real military operations for an extended period of time. Even the "novices" had a basic understanding of map graphics and theoretical knowledge of how battles are fought. A true novice to military operations would not know the meaning of phase lines, control measures, avenues of advance, or of symbols indicating unit size and composition. The operations order and the command net messages would be largely uninterpretable by them.

The cognitive abilities tests were somewhat disappointing, with the exception of the Building Memory test. The very high average on this test suggests that there is a relationship between the abilities measured by the test and the requirements of Army command and staff assignments. Since much of Army staff work is map-based, this may reflect extra experience with map interpretation. Alternatively, it may reflect the operation of a successful personnel assignment system, i.e., a system that usually picks officers who have the capabilities required for successful performance. A third possibility is self-selection, i.e., officers who are comfortable with maps are more likely to select a career requiring their use. Two features of the data are relevant to these

explanations. First, experience is not the only explanation. The six perfect scores included officers with two and three years experience, while the seven scores below the mid point of the scale included officers with 15 and 19 years of service. Second, despite the restricted range of scores, there was a significant association between Building Memory and visualization accuracy. If visualization as measured in this experiment is strongly related to job performance in a staff or command assignment, then the data support the idea that appropriate skills are rewarded in assigning officers to jobs. This appears to be a fruitful area for future work. While memory for maps and spatial relationships have been extensively studied, especially in military research (Thorndyke, 1980), they have not been previously related to individual performance requirements in command and control.

The other four cognitive abilities tests are not convincingly related to our accuracy measures. The correlations between advanced vocabulary scores and both location measures may be due to the fact that vocabulary scores are related to age and age is related to the location measures. The paper folding test scores are related to both strength measures, but in opposite directions. There is no obvious explanation of how the ability to visualize the results of folding a diagram could increase the accuracy of strength monitoring and decrease the accuracy of strength predictions.

The cognitive testing session did not induce the enthusiastic cooperation of all of the participants, judging from some of the feedback received by the experimenters. The tactical scenario session, on the other hand, was seen by most participants as obviously relevant to their professional concerns. Thus, uncorrelated variance could have been introduced in the cognitive measures through differences in motivation thus reducing the correlations of these measures with the location and strength judgments.

An empirical generalization about what makes a task hard was developed to account for the differences between scenarios and tasks as they relate to the iteration factor. It appears that much of the data could be accounted for by the assertion that errors tend to increase when the situation changes rapidly. This holds both for predicting future states and for visualizing the current state. It applies to both location and strength judgments and potentially for other variables that could be tracked, such as consumption rates for fuel and ammunition.

To relate this rule to the pattern of results displayed in Figures 1-4, it is necessary to introduce some facts about deliberate attack scenarios. Initially, the forces are widely separated. The attacking side tries to move rapidly to assault positions and begin the attack on an isolated portion of the defensive force before they have time to reinforce the position. As the attacker nears his objective, he usually must overcome defensive barriers, such as minefields and tank ditches. Once the forces come within range to use their direct fire weapons systems (tank main guns, anti-tank missiles, etc.), movement slows. A standard rule applied by many of our participants was one minute per kilometer for cross-country movement and three minutes per kilometer once the forces are engaged. Significant delays often occur at barriers.

One would expect more movement, hence larger location errors, early in a deliberate attack and less movement, smaller errors in locations, later in a battle. Changes in strength of the forces would have the opposite pattern. Until the forces are within direct fire range, only sporadic losses

due to artillery and aircraft would be expected, with perhaps some losses to direct fire engagements between the attacker and scouts. Casualties would mount quickly once the main battle commenced. Thus, our data should show more accurate strength judgments early in a battle and more error later. With some minor anomalies, this is the pattern shown for both strength and location judgments in Figures 1-4.

This rule can also account for the differences in accuracy between the scenarios. The armor task force moved more rapidly and had its strength depleted more rapidly than the mechanized infantry task force. Therefore, it should have and did result in less accurate judgments when error is measured in kilometers and number of vehicles. (The previous assertion that the armor scenario was more predictable is still true on a relative basis. When comparing the amount of error to the amount of movement that actually occurred, the average error in the armor scenario was typically around one-half to one-third of the distance moved, whereas, in the infantry scenario, the errors were on the order of two to three times the distance moved on the first iteration of judgments. In both scenarios, the participants adapted to the rate of progress being achieved by the unit, but still obtained better relative accuracy in the armor scenario on subsequent iterations.)

This generalization leads to some general expectations for the pattern of results that should occur when applying the measurement technique to conditions not previously tested. For example, a defend-in-sector mission is characterized by very little movement by the defensive side and few casualties prior to the main assault. A covering force operation would involve much more movement and a higher casualty rate than a static defense. A night attack is characterized by slower movement over shorter distances than a day attack. Thus location of forces should be more predictable in a night attack, contrary to what one might assume, knowing how notoriously difficult such operations are to control. (Note that this prediction is based upon our measurement procedure which supplies accurate starting locations. A frequent problem in night operations is that people do not know where they are.)

Analysis of the experimental tasks suggests that no tidy theoretical account of performance based upon the use of an underlying schema, a strategy for using an episodic model, or an integrated set of constraints is likely to successfully describe what the participants did in this experiment. For the two prediction tasks, the facts given previously about deliberate attack missions indicate that there is a logical connection between correctly estimating where forces will be located and knowing when direct fire casualties will occur. For the two visualization tasks, the connection is more tenuous: If you failed to hear or to remember the reports of combat losses, knowledge of where units are located may give some advantage in estimating what the losses should have been. Similarly, if you know where the enemy is (the defenders don't move much), and hear which of your own elements are taking casualties, you may be able to deduce your own locations.

The two location judgment tasks are fairly similar, though the time scale may differ. The radio messages give only sporadic information on friendly force locations. To know where all of your elements are at a given movement, you have to deduce the locations based upon where they were the last time you heard, where they are supposed to be going, and what kind of progress you would expect them to make in the interim. In terms of absolute error, there should be an

advantage to visualizing locations over predicting locations, since more recent location information is usually available and since deviations from the plan forced by the opponent or directed by the commander are heard on the command net. Figures 1 and 2 indicate that short range predictions are as accurate, on the average, as judgments of current locations, but that visualization is generally more accurate than longer range predictions.

The two strength judgment tasks do not have as much in common. The strength visualization task involves remembering what you heard and guessing how much did not get reported -- in particular, what is happening to the enemy side. Predicting strength, on the other hand, involves a number of imponderables, such as what the influence of aircraft and artillery will be, when the forces will close, how accurate the gunnery will be, whether the attacker will be adequately screened by smoke, how long it will take to breach defensive barriers, and whether the defensive position will be reinforced in time to halt the attack. In fact, the strength prediction task may be so difficult that no serious attempt is made to consider and evaluate all of these factors. A participant may instead rely upon previous experience or some form of guesswork to supply the numbers asked for by the experimenter. The correlations in Table 4, which show a significant association only between predicting locations and visualizing locations, are consistent with the similarities among tasks as just discussed.

The low association among our performance measures does not argue against the existence of a coherent internal representation of battlefield events as discussed by the cognitive theorists, but does suggest that it will be difficult to probe such structures using global quantitative measures of the type we obtained in this experiment. The verbal protocols, especially those predicting future actions by the commander and by the opposing force, may provide data more relevant to the theoretical issues.

The transition between verbally expressed understanding of a complex environment and task performance within that environment is a link that will be difficult to establish, since there often exists a short cut method to achieve satisfactory answers without indulging in complex reasoning. Furthermore, individual differences in cognitive structures may confer differential advantages upon differing sorts of tasks and individual differences in using cognitive structures to perform tasks may also contribute differentially to performance scores on different tasks.

References

- Beach, L. R. (1990). Image theory: Decision making in personal and organizational contexts. Chichester: Wiley & Sons.
- Bolger, D. P. (1986). Dragons at war. New York: Ivy Books.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive test. Princeton, NJ: Educational Testing Service.
- Johnson-Laird, P. N. (1983). Mental Models: toward a cognitive science of language, inference, and consciousness. Cambridge, MA: Harvard Press.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgement under uncertainty: Heuristics and biases. New York: Cambridge Press.
- Myers, I. B. & McCaulley, M. H. (1988). Manual: S guide to the development and use of the Myers-Briggs type indicator. Palo Alto, CA: Consulting Psychologists Press.
- Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals and understanding. Hillsdale, NJ: Erlbaum.
- Thorndyke, P. W. (1980). Performance models for spatial and locational cognition. The Rand Corporation, R-2627-ONR.
- Thuring, M., & Jungerman, H. (1986). Constructing and running mental models for inferences about the future. In B. Brehmer, H. Jungerman, P. Lourens, and G. Sevon (Eds.), New directions in research on decisionmaking. Elsevier Science Publishers V. B. (North-Holland).

BATTLE FORECASTING
AND VISUALIZATION

ANSWER BOOK

Scenario #6



PLEASE

DO NOT TURN THIS PAGE

UNTIL ASKED TO DO SO

INSTRUCTIONS

Using the tape recorder which has been provided for you, please give your assessment of the Plan you have just read. Include in your assessment a discussion of the Task Organization, Scheme of Maneuver, and use of mass, space and time factors.

Additionally, please estimate on a scale of 0% to 100% the probability of taking Objective Spade.

PLEASE

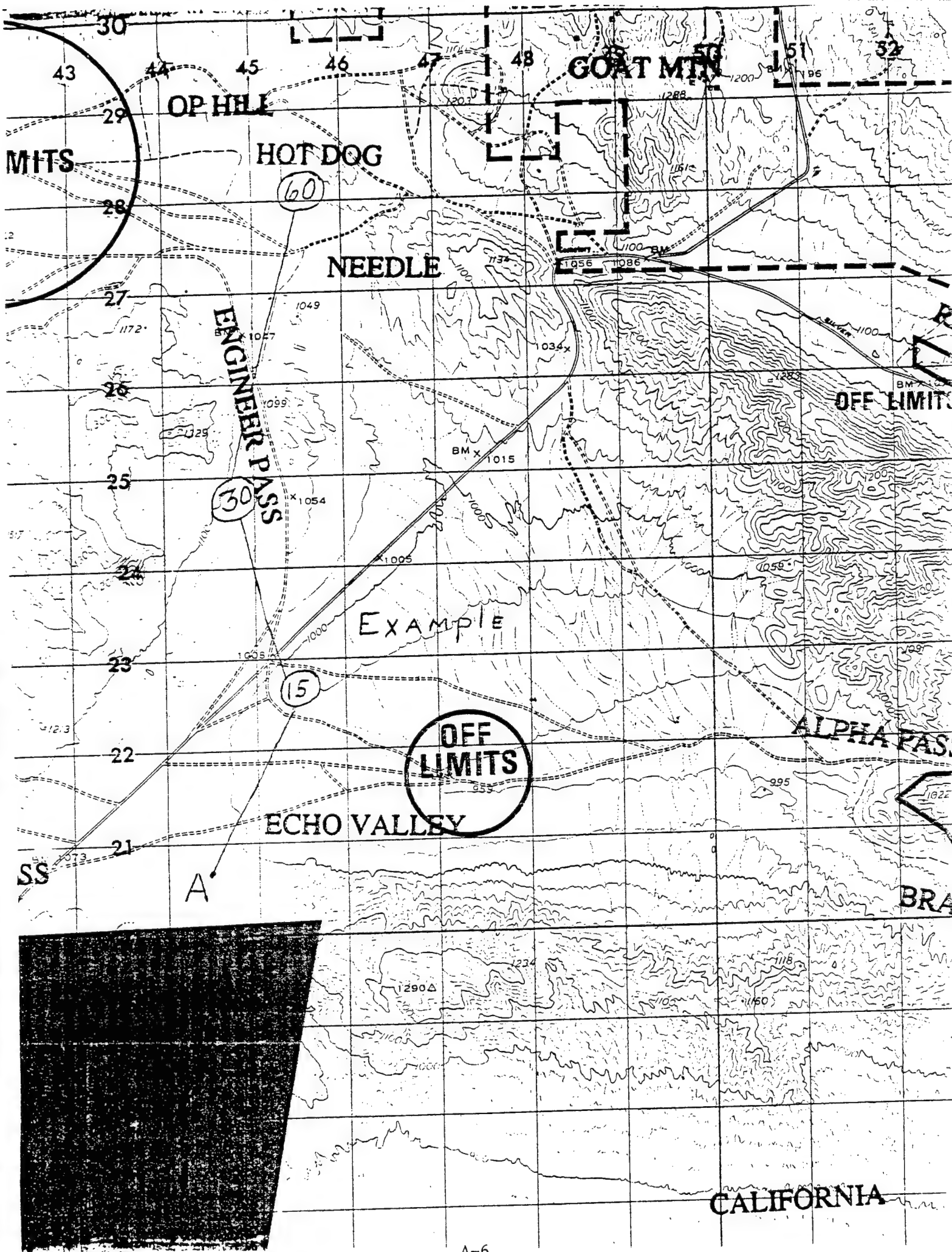
DO NOT TURN THIS PAGE

UNTIL ASKED TO DO SO

INSTRUCTIONS

Discuss what you think will happen in the battle during the next hour.

Please turn to the next page of the Answer Book and give detailed predictions for movement and attrition.



Please provide predictions of future strength totals by filling in the chart below. Indicate how confident you are in your prediction by providing a range around your best guess. For example if you thought there would be around 20 tanks remaining but the total could be off by 4 on either side, you should respond 20 ± 4 .

	Current Time:0835	Future (+15) Time:0850	Future (+30) Time:0905	Future (+60) Time:0935
Friendly Tanks	30	30 ± 2	24 ± 2	18 ± 2
APCS	15	15 ± 2	12 ± 1	10 ± 1
OPFOR Tanks	26	26 ± 2	22 ± 1	20 ± 1
APCS	46	26 ± 2	25 ± 2	22 ± 2

PLEASE

DO NOT TURN THIS PAGE

UNTIL ASKED TO DO SO

INSTRUCTIONS

Discuss what you think will happen in the battle during the next hour.

Please turn to the next page of the Answer Book and give detailed predictions for movement and attrition.

Please provide predictions of future strength totals by filling in the chart below. Indicate how confident you are in your prediction by providing a range around your best guess. For example if you thought there would be around 20 tanks remaining but the total could be off by 4 on either side, you should respond 20 ± 4 .

	Current Time:0850	Future (+15) Time:0905	Future (+30) Time:0915	Future (+60) Time:0950
Friendly Tanks	28	24 ± 2	20 ± 2	17 ± 2
APCS	12	10 ± 2	5 ± 1	4 ± 1
OPFOR Tanks	26	24 ± 1	21 ± 1	18 ± 2
APCS	46	42 ± 2	38 ± 2	32 ± 2

PLEASE

DO NOT TURN THIS PAGE

UNTIL ASKED TO DO SO

INSTRUCTIONS

Please discuss what happened during the previous portion of the battle in terms of attrition and activities.

Now turn to the next page and plot on the map where you think the lead platoon of each team is located. Use the symbols A, B, C, and D for labeling.

Please provide estimates of current strength by completing the chart below. Indicate how confident you are in your estimates by providing a range around your estimate. For example, if you estimate there are 15 tanks currently, but the total could be off by 4 on either side, you would respond 15 \pm 4.

		<u>Current Estimate</u>	<u>Range</u>
Friendly	Tanks	6	± 2
	APCs	6	± 1
OPFOR	Tanks	22	± 1
	APCs	40	± 2

PLEASE

DO NOT TURN THIS PAGE

UNTIL ASKED TO DO SO

INSTRUCTIONS

Discuss what you think will happen in the battle during the next hour.

Please turn to the next page of the Answer Book and give detailed predictions for movement and attrition.

Please provide predictions of future strength totals by filling in the chart below. Indicate how confident you are in your prediction by providing a range around your best guess. For example if you thought there would be around 20 tanks remaining but the total could be off by 4 on either side, you should respond 20 ± 4 .

	Current Time:0905	Future (+15) Time:0920	Future (+30) Time:0935	Future (+60) Time:1005
Friendly Tanks	26	22 ± 1	18 ± 2	15 ± 2
APCS	11	9 ± 1	6 ± 1	4 ± 1
OPFOR Tanks	26	25 ± 1	23 ± 1	21 ± 1
APCS	46	44 ± 1	41 ± 2	41 ± 1

PLEASE

DO NOT TURN THIS PAGE

UNTIL ASKED TO DO SO

INSTRUCTIONS

Please discuss what happened during the previous portion of the battle in terms of attrition and activities.

Now turn to the next page and plot on the map where you think the lead platoon of each team is located. Use the symbols A, B, C, and D for labeling.

UCKY FUSE
IMPACT

W-3 CC

PLACON
DO

HILL 780

HILL 760

HILL 699

HOO HILL

HILL 824

HILL 781

HILL 720

LEY

FINGER

JOHN WAYNE HILLS

TAC 501

PORTA HILL W

HILL 758

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

HILL 720

Please provide estimates of current strength by completing the chart below. Indicate how confident you are in your estimates by providing a range around your estimate. For example, if you estimate there are 15 tanks currently, but the total could be off by 4 on either side, you would respond 15 \pm 4.

		<u>Current Estimate</u>	<u>Range</u>
Friendly	Tanks	21	± 2
	APCs	9	± 1
OPFOR	Tanks	24	± 2
	APCs	45	± 1

PLEASE

DO NOT TURN THIS PAGE

UNTIL ASKED TO DO SO

INSTRUCTIONS

Please discuss what you expect to happen in the remainder of the battle in terms of attrition and activities.

Additionally, please estimate on a scale of 0% to 100% the probability of taking Objective Spade.

PLEASE

DO NOT TURN THIS PAGE

UNTIL ASKED TO DO SO

INSTRUCTIONS

Please discuss what happened during the previous portion of the battle in terms of attrition and activities.

Now turn to the next page and plot on the map where you think the lead platoon of each team is located. Use the symbols A, B, C, and D for labeling.

Please provide estimates of current strength by completing the chart below. Indicate how confident you are in your estimates by providing a range around your estimate. For example, if you estimate there are 15 tanks currently, but the total could be off by 4 on either side, you would respond 15 \pm 4.

		<u>Current Estimate</u>	<u>Range</u>
Friendly	Tanks		3 \pm 1
	APCs		5 \pm 2
OPFOR	Tanks		20 \pm 2
	APCs		40 \pm 3

APPENDIX B

BATTLEFIELD VISUALIZATION PROTOCOL SCENARIO #4

1. Introduction: Mission/Situation/Plan

a. "What we will be asking you to do is, first, familiarize yourself with the plan, then give us some predictions and situation assessments."

b. "The first three or four pages of the flip chart give information about the plan.
[SHOW FLIP CHART]

We also have a map and ops overlay. Some information you want may be missing."

[SHOW MAP AND OVERLAYS]

c. "We(I) should point out that this scenario was selected at random from the NTC Database, that is, not because of particularly good or poor performance or because it is typical but largely because there was enough information for us to use it for this purpose."

d. "Initially, we(I) would like for you to assess the plan in terms of the mission, the friendly situation, and the enemy situation. Then give us an estimate in terms of a number between zero and 100 of how successful you think the Task Force will be at taking Objective 32."

e. [AS SUBJECT GETS INTO PROBLEM, BE PREPARED TO REITERATE WHAT WE ARE ASKING FOR, MENTION THAT MORE INFO WILL BE GIVEN IN MAP AND TAPE UPDATES. ANSWER SIMPLE QUESTIONS.]

f. ASK FOR CRITIQUE OF PLAN.

***GET: Original Estimate of Likelihood of Success

2. Show Overlay @ 06:15 [Infiltration, Traffic Jam]

Explain Color Code

PREDICTIONS 1 [15 30 60] Example, Verbal, Map, Strength

3. Update Tape - Segment 1 (0615-0630)

VISUALIZE current situation: Verbal, Map, Strength

4. Show Overlay @ 06:30

PREDICTIONS 2 [15 30 60] Verbal, Map, Strength

5. Update Tape - Segment 2 (0630-0700)

VISUALIZE current situation: Summary, Map, Strength

6. Show Overlay @ 07:00

PREDICTIONS 3 [15 30 60] Verbal, Map, Strength

7. Update Tape - Segment 3 (0700-0800)

VISUALIZE current situation: Summary, Map, Strength

8. Show Overlay @ 08:00

Ask for what happened the remainder of the battle.

***GET: Final Estimate of Likelihood of Success

9. WRAP-UP: Final situation

Ask for critique of information Available and Procedure

BATTLEFIELD VISUALIZATION PROTOCOL SCENARIO #6

1. Introduction: Mission/Situation/Plan

a. "What we will be asking you to do is, first, familiarize yourself with the plan, then give us some predictions and situation assessments."

b. "The first three or four pages of the flip chart give information about the plan.
[SHOW FLIP CHART]

We also have a map and ops overlay along with Intel and Fire Support overlays. Some information you want may be mission."
[SHOW MAP AND OVERLAYS]

c. "We(I) should point out that this scenario was selected at random from the NTC Database, that is not because of particularly good or poor performance or because it is typical but largely because there was enough information for us to use it for this purpose."

d. "Initially, we(I) would like for you to assess the plan in terms of the mission, the friendly situation, and the enemy situation. Then give us an estimate in terms of a number between zero and 100 of how successful you think the Task Force will be at taking Obj SPADE."

e. [AS SUBJECT GETS INTO PROBLEM, BE PREPARED TO REITERATE WHAT WE ARE ASKING FOR, MENTION THAT MORE INFO WILL BE GIVEN IN MAP AND TAPE UPDATES. ANSWER SIMPLE QUESTIONS.]

f. ASK FOR CRITIQUE OF PLAN.

***GET: Original Estimate of Likelihood of Success

2. Show Overlay @ 08:35

Explain Color Code

PREDICTIONS 1 [15 30 60] Example, Verbal, Map, Strength

Show Overlay @ 08:50. Ask for nothing.

3. Update Tape - Segment 1 (0850-0905)

VISUALIZE current situation: Verbal, Map, Strength

4. Show Overlay @ 09:05

PREDICTIONS 2 [15 30 50]

5. Update Tape - Segment 2 (0905-0935)

VISUALIZE current situation: Summary, Map, Strength

6. Show Overlay @ 09:35

PREDICTIONS 3 [15 30 60] Verbal, Map, Strength

7. Update Tape - Segment (0935-1035

VISUALIZE current situation: Summary, Map, Strength

8. Show Overlay @ 10:35

Ask for what happened the remainder of the battle

***GET: Final Estimate of Likelihood of Success

9. WRAP-UP: Final situation

Ask for critique of Information Available and Procedure

BATTLEFIELD VISUALIZATION ADDITIONAL INFO SCENARIO #6

Plan Evaluation

Locations of Spade subdivisions 11, 12, 32 unavailable
Checkpoint 40 is at PL Aggie, N. edge of Dreadnought

Situation @ 08:35

10 minutes prior to LD, TF lost 2 Scouts vic 2 Red guys down on circle below Dreadnought.
When Tm C was a few meters short of LD, they were hit by ARTY and took a few losses.
Status board shows strength already down from initial level. They already had a few systems down from maintenance and did not start at full strength.

Situation @ 08:50

Point out comparison of strength chart to their first prediction
Among losses was Tm C Cdr, Co XO took charge
TF reacted to ARTY by switching to an alternate route that goes a bit N. of Dreadnought
(Tm A is up there)
re TAPE: "Bone" is at checkpoint 40